

Involving linguists in developing AI for language learning

Kacper Łodzickowski (Vice President, AI Learning Capabilities, Pearson)

This talk will provide practical ideas for how linguists can be involved in developing commercial artificial intelligence (AI) systems for second language (L2) learning.

First, we will inspect two types of AI-based L2 capabilities: automated assessment of complex performances and automated creation of educational content.

Then, we will discuss how leveraging linguistic expertise can help inform the three key stages of researching & developing such capabilities, namely:

1. **System design**, which involves determining learning outcomes and relevant assessment constructs as well as collaborating with engineers on selecting meaningful computational features.
2. **Data collection**, which requires an appreciation of the spectrum of target language varieties and L2 learner performances, so that the data used for model training is reasonably free from latent biases.
3. **System validation**, which relies on a human-in-the-loop approach for interpreting the practical significance of model training results and monitoring models in a production environment (where their performance may change unexpectedly as a function of unpredictable customer behaviour).

Finally, we will evaluate an industry perspective on how higher education institutions could expand their linguistics programmes to train the next generation of graduates ready to partner with businesses and maintain professional standards for linguistics in commercial AI applications.